



RIETI Discussion Paper Series 11-J-002

日本企業の研究開発活動から商業化へのラグ構造の分析

鈴木 潤

政策研究大学院大学



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所

<http://www.rieti.go.jp/jp/>

日本企業の研究開発活動から商業化へのラグ構造の分析

鈴木 潤（政策研究大学院大学）

要 旨

RIETI 発明者サーベイのデータと、特許データベースから得られる統計データを用いて、研究開発のタイムラグの構造を調べた。研究開発活動に対するインプットからアウトプット、アウトカムまでにいたるタイムラグは、いくつかの性格の異なる期間が合成された結果であると考えられる。本分析では、新たな研究開発プロジェクトを開始してから特許を出願するまでのラグと、特許を出願した後に利用開始に至るまでのラグを調べた。これらを合計した期間の長さは左に偏った分布形状を持っており、中央値が約 33 ヶ月であった。特許出願後にその特許を利用するか否かに影響を与える因子と、利用開始までの期間に影響を与える因子は独立性が高く、利用の意思決定は主として発明内容が技術的に高い価値を持つかどうかにかき依存する。一方、利用開始までの期間は、研究プロジェクトの規模や企業戦略により大きな影響を受けることがわかった。また、近年、研究開発期間の短縮や加速化などが生じているとする意見を支持するような確実な証拠は得られなかったが、特許制度の改正が、出願後利用開始までの期間に影響を与えている可能性があることがわかった。

キーワード： 研究開発、特許出願、商業利用、タイムラグ、発明者サーベイ、特許統計

JEL Classification: 031, 034

RIETI ディスカッション・ペーパーは、専門論文の形式でまとめられた研究成果を公開し、活発な議論を喚起することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、（独）経済産業研究所としての見解を示すものではありません。

本稿は、経済産業研究所における「日本企業の研究開発の構造的特徴と今後の課題」の研究成果の一部である。RIETI 発明者サーベイ・データの分析に参加するという貴重な機会ならびに研究内容への多くの助言と示唆をいただいた長岡貞男研究主幹に深く感謝する。また、大湾秀夫教授（東京大学）をはじめとする研究会メンバーや、後藤晃教授（政策研究大学院大学）、及川耕造理事長、藤田昌久所長、富田秀昭研究コーディネーターをはじめとする経済産業研究所の関係者の方々から、ワークショップや研究会の場で多くの貴重なご意見や示唆を頂いた。

1. はじめに

現在、科学技術に対する公的な支援は、世界的に冷戦時代以来の歴史的な高レベルに達し、なおかつ急速に増大しつつある。日本あるいは世界経済の将来を展望するならば、様々な分野や局面においてイノベーションが今後ますます重要なものとなっていくであろうという点に異を唱える人は少ないだろう。しかし、科学技術政策や産業政策、経済政策などがイノベーションとどのような関係を持っているのか、またそれらをより効果的に結びつけるためにはどのようなことが必要なのか、これらはきわめて今日的なリサーチ・クエッションであると同時に、過去 50 年以上にわたって様々な研究者が取り組んできた古典的課題でもある。

基礎的な研究開発活動には、公的な支援が必要であるとされている。しかしながら、基礎的な研究活動が最終的に経済的なリターンに結実するまでの経路はきわめて複雑で長い時間を要する場合が多く、研究プロジェクトから直接的に生じる新製品や新サービスのみに焦点を当てた費用便益分析は不十分であり、誤った解釈を生じさせる恐れが大きいことが指摘されている (Mowery, 1994 など)。基礎的な研究に対する支援についてはこのような議論がある一方で、より応用的な研究開発活動に対する国の関与や支援を容認する、あるいは求めるような社会的動きがあることも事実である。米国のブッシュ前大統領は、米国産業の“競争力”を強化するために、科学者や技術者の養成と自然科学研究支援に対する政府支出の増額を提案した。また、オバマ大統領が不況脱出の切り札の一つとして、研究開発減税の恒久化を再度提案したことは記憶に新しい。ただし、公的活動・企業活動によらず研究開発とイノベーションの関係の大部分は依然としてブラックボックスであり、厳密には研究開発活動はイノベーションの十分条件ではないし、必要条件であるとも言えない。すなわち、研究開発を行えば必ずイノベーションに結実するわけではなく、また研究開発なしにイノベーションを実現できないわけでもない。一般的に、持続的なイノベーションの実現において、研究開発活動が極めて重要な役割を担っているということは産業界や学界を問わず共有された認識であろう。ただ、これを具体的なデータに基づいて証明することは、かなり難しいことなのである。

市場に近い研究開発でさえ、インプットとアウトプットの分析が難しい理由の一つに、研究開発活動にリソースが投入された後にその成果が観察されるようになるまでにある程度の時間がかかるという、タイムラグの存在があげられる。例えば、企業の経営者は研究開発投資の成果をいつ判断すればよいのだろうか？企業に対する公的な研究開発補助金の効果や研究開発優遇税制の効果は、いつごろどのように評価すればよいのであろうか。研究開発の短期成果志向が進んでいるという指摘も聞かれるが、そのような現象は具体的に実証可能であろうか。最近では国民経済計算における無形資産（研究開発の成果は無形資産の全てではないが、その重要な要素として扱われる）の推計においても、研究開発のタイムラグが検討課題のひとつとして指摘されている (Corrado et.al., 2006、元橋, 2009 など)。

本研究は、これまであまり深い分析がなされてこなかった、研究開発活動とその成果としての特許、そしてその特許が商業的利用に至るまでの一連のタイムラグの構成と、それらに影響を与える因子を解明しようとする試みである。

2. 研究開発のタイムラグに関する先行研究

研究開発に関する経済分析の泰斗である Griliches は、恒久棚卸法の考えを援用して、研究開発活動への投資と技術的な知識ストック¹の関係を以下のように定式化した (Griliches, 1984) :

$$K_t = \sum (1 - \delta)^i R_{t-i}$$

K_t : t 期における知識ストック量

R_{t-i} : 過去の研究開発投資

$(1 - \delta)^i$: ラグ分布と知識の減衰の影響

$(1 - \delta)^i$ は成果結実までのタイムラグ分布と、過去の知識の減衰の影響の両方により定まるパラメータである。しかし、その当時はそれらの影響を推定するための精度の高いデータが入手できなかった。また、何種類かのラグ分布と初期値についてシミュレーションした結果がそれほど大きな影響を与えなかったため、Griliches は $\delta = 0.15$ を暫定的に用いることを提案している。

日本におけるこの種の分析の草分けとなった、後藤ほか, 1986 の分析では、知識の減衰係数と平均ラグが用いられた。すなわち :

$$K_t = (1 - \delta) K_{t-1} + R_{t-\theta}$$

K_t : t 期における知識ストック量

δ : 知識の減衰係数

$R_{t-\theta}$: θ (=研究開発平均ラグ) 期前の研究開発活動

δ と θ は、技術分野や組織 (企業)、時期その他の要因で定まるパラメータである。現実には研究開発活動の成果はある程度の時間経過の後に徐々に分布ラグを伴って結実するものと考えられるが、平均ラグによる定式化は過去に実施された研究開発活動の成果が全て θ の時間経過を経て一気に結実するという仮定を置いている事になる。これは、研究開発成果の分布ラグについての適切な情報を得ることが難しいため、やむなく採られた方法であったが、その後の分析でも近年に至るまで基本的に同様のアプローチが繰り返されてきた。後藤らは、経済企画庁が 1982 年に実施した「企業行動に関するアンケート調査」の結果を元に、研究開発から企業化までに要する平均ラグ期間を求めている。産業別に計算された平均ラグ期間は、無機化学で 2.9 年、医薬品で 4.6 年、通信機器で 2.0 年、自動車で 2.8 年などであった。

¹ 組織に蓄積される知識としては、例えば販売や流通のノウハウなど技術的知識以外にも様々な種類と源があり、研究開発活動の結果として得られる知識はその一部である。ここで用いられているモデルでは、研究開発活動への投資は新たな技術的知識を得るためのインプットであり、蓄積された知識を元にして発明や新製品・プロセスなどのアウトプットが生み出されることが想定されている。もちろん、研究開発投資額や特許数などは不完全な代理指標であるし、このようなある種のリニアモデルに対する批判や、また物的資産とは異なり原則的には経時劣化することのない知識という資産に対して有形資産と同様の恒久棚卸法を適用することの是非も議論されている。しかし、利用可能なデータの制約等から、これ以上に説得力のあるモデルもいまだ登場しておらず、多くの経済分析で利用される古典的モデルとなっている。

文部科学省科学技術政策研究所は、1999 年に「研究開発関連政策が及ぼす経済効果の定量的評価手法に関する調査」として、資本金規模上位の企業を対象にサーベイを実施した（科学技術政策研究所、1999 年）。調査の対象は企業の個別研究開発プロジェクトであり、630 の有効回答を得ている。この調査で調べられている研究開発ラグ関連の項目は、“研究開発に要した期間”と、“研究開発終了後、市場に導入されるまでの期間”である。これらの質問項目はいずれも 6 カテゴリーの区間で回答を求めている。研究開発に要した期間については、全体では「2 年以上～3 年未満」が最も多く 27.2%であった。基礎研究ほど研究期間が長く、開発研究ほど短くなっている。各カテゴリーの中間値を代表値とし、加重平均で求められた平均は 3.6 年である。また、その研究開発が終了した後、その技術を用いた製品等が市場に導入されるまでに要した期間について、上記と同様の計算で求められた平均は 1.2 年である。このレポートでは、これらをさらに合成し、研究開発終了後市場に導入されるまでの期間に研究開発期間の 1/2 を加えて、“研究開発投資のタイム・ラグ”と呼んでいる。すなわち、 $3.6 \div 2 + 1.2 = 3.0$ 年が全体の平均である。またこのレポートでは使用した研究費と研究開発タイム・ラグのクロス集計を行い、使用金額の多い研究開発成果ほど、市場に出るまでの期間が長いという結論を得ている。なお、この調査では研究開発の成果としての技術の特許出願したかどうかについても尋ねている。それによると、全体のおよそ 60%のプロジェクトで特許が出願されており、1 プロジェクトあたりの平均出願件数は 20.7 件であった。研究が開始された段階別に見ると、基礎研究段階から始められた研究開発ほど出願されている割合が高い結果となっている。

経済産業省は、2004 年に「研究開発促進税制の経済波及効果に係る調査」として、研究開発支出や売上の大きな企業 1313 社を対象にサーベイを実施した（横山・吉本、2005 年）。調査対象は各企業の主要業種部門における研究開発の事例で、有効回答は 222 社の 340 事例である。質問項目や回答方法などは前述の科学技術政策研究所の調査をほぼ踏襲しており、“研究開発に要した期間”と“研究開発終了後から市場に投入されるまでの期間”を 8 ないし 5 カテゴリーの区間で尋ねている。これらの回答を元に、前述の調査と同様カテゴリーの中間値を代表値とみなす計算方法によって求められた平均は、研究開発期間が 4.1 年、研究開発終了後市場導入までが 1.6 年である。そして、研究開発終了後市場に導入されるまでの期間に研究開発期間の 1/2 を加えた“研究開発タイムラグ”は、このレポートでは 3.7 年が全体の平均値となっている（中央値は 3.0 年）。産業別に計算された総平均ラグ期間は、食品工業で 3.1 年、化学工業（医薬品含む）で 4.2 年、電気機械工業で 3.3 年、輸送用機械工業で 3.4 年などであった。

科学技術政策研究所は、上述と同様の調査を 2009 年にも実施している（科学技術政策研究所、2010 年）。調査の方法や設問は、1999 年の調査とほぼ同じであるが、回答の選択肢としてより細かい 7 カテゴリーを採用している。上述と同様の計算方法に従って得られた研究開発期間の平均値は 42.7 ヶ月、研究開発終了時から市場導入までに要した期間は 20.0 ヶ月であり、“研究開発投資のタイム・ラグ”は 41.4 ヶ月（3.5 年）であった。

3. 研究開発のタイムラグの構成要素

第2節で述べたように、研究開発活動の平均タイムラグは、時代により産業により違いはあるものの、ほぼ2年～4年程度とする調査結果が多い。これらの調査で調べられているのは、実際には“研究開発に要した期間”と“研究開発終了後に市場に導入されるまでの期間”であった。しかし、このようなラグの捕らえ方は果たして適切なものであろうか？

(1) アウトプット・ラグに関する仮説

研究開発のインプットとアウトプット／アウトカムの実証分析に用いられる代表的な指標としては、研究開発投資額、研究開発人材（以上、インプット）、科学論文や特許（以上、アウトプット）、新製品や新サービス(の上市や売上高)、付加価値額や利益、企業価値(以上、アウトカム)などが一般的である。これらのうち、研究開発活動により生まれた知識としての成果が科学論文や特許として結実するまでのタイムラグは、研究開発費が投入された（研究開発プロジェクトが開始された）時点から、論文や特許が投稿あるいは出願された時点までの期間をタイムラグと捉えるべきであろう。これは、研究開発のアウトプット・ラグに相当する。産業界では論文よりも特許のほうが圧倒的に重要であるから、これ以降は特許を中心として検討を進める。日本を始めとして多くの国で採用されている先願主義の *winner takes all* の競争環境では、研究開発により重要な知見が得られた場合にはなるべく早期に特許を出願する必要があるから、特許が出願される時点は必ずしも研究開発が終了した後になるとは限らない。すなわち、“研究開発に要した時間”からアウトプット・ラグを推定しようとするよりも、“研究開始から特許出願までに要した時間”を調べるほうがより適切であると考えられる。それでは、このアウトプット・ラグはどのような因子から影響を受けるだろうか。

研究プロジェクトの規模が、アウトプット・ラグに影響を及ぼすことは十分に考えられる。大きなリソースを投入して困難な課題に挑戦し、結果として多数の特許が出願されるような大規模研究開発プロジェクトでは、一定の成果が得られるまでには長い時間がかかるのではないだろうか。一方、同じような研究課題に多人数で取り組む場合には、少人数で取り組む場合よりも早く成果が得られるであろう。また、研究プロジェクトに外部の機関がかかわっている場合には、権利持分の交渉などで単独のプロジェクトよりも出願までに時間がかかる可能性があるだろう。

科学技術政策研究所の調査でも明らかにされているように、研究プロジェクトの性格として、基礎的な内容を含むものかどうかにより特許出願が影響を受けることが考えられる。

また、発明は必ずしも狙った結果として得られるとは限らず、研究プロジェクトの開始当初には思いもかけなかった副次的な産物として生まれることもあるし、特定の研究開発とは関係の無い場面やセクションで生まれることもあるだろう。発明が発明を生むような先行発明に依拠した特許の場合には、当初の研究プロジェクト開始からは長い時間が経っているかもしれない。そのような発明の経緯も、ラグに影響を与える可能性がある。

そして、発明そのものの性格、例えばモノの発明か方法の発明か、またどの程度画期的な内容を含んでいるかなどが、企業の特許出願行動やラグに影響を与えることが考えられる。

最後に、より一般的な環境条件、例えば規模の大きな企業ではしっかりした知財部のサポートがあり、速やかな出願が可能になるかもしれない。また、技術分野の性格により研究開発そのものにかかる時間は大きく異なるだろう。先行研究が示唆しているように、技術開発競争や市場競争の激化により、最近になるほど早期に結果が求められるようになってきたという傾向が見られるかもしれない。

すなわち、本分析では LRP_{it} (i 分野・ t 期における R&D のアウトプット・ラグ) が R&D プロジェクトの規模 ($RDsize_{it}$)、R&D プロジェクトの性格 ($RDchar_{it}$)、発明のプロセス (INV_{it})、特許の技術的な画期性 ($TVAL_{it}$)、企業の規模 ($FIRM_{it}$) の関数であると仮定する。

$$LRP_{it} = f(RDsize_{it}, RDchar_{it}, INV_{it}, VAL_{it}, FIRM_{it})$$

(2) アウトカム・ラグに関する仮説

研究開発のアウトプット・ラグについては上述のような仮説に基づく変数を利用して分析を行うが、一方、特許を出願した後でその技術に基づく新製品や新サービスが市場に導入されるまでには、通常はさらに一定の時間が必要である。研究開発のアウトプットからアウトカムに至るまでのラグである。ただし、特許行政年次報告書, 2010 年などから明らかなように、日本では出願された特許のうち審査請求されるものは約 2/3 しかないし、最終的に登録されるものは半分程度である。また登録された特許に関しても、その全てが新規の製品やサービス、プロセス、ライセンスなどに利用されるわけではない（“利用” の定義にもよる）。本分析では、特許出願後いずれかの時点でその技術を自社で利用するかどうかの意思決定が行われることを仮定する。そして、“利用されるかどうか” と “利用されるまでの時間” には、それぞれ異なる因子が影響を与えていることが予想される。ただし、アウトカム・ラグに関するデータの形態は典型的な打ち切り（Censored）データであって、“利用されるかどうか” と “利用されるまでの時間” に影響を与える因子の完全な独立性を仮定できないので、サンプルのセレクション・バイアスを考慮に入れた分析を行う必要がある。

出願された特許が審査請求や登録にまで至るかどうか、あるいは実際に利用されるかどうかは、その技術の開発にどれだけのリソースが投入されたか、すなわちサンク・コストにある程度依存するものと考えられる。また、その成果が当初の目標どおりのものであったかどうか、すなわち発明の経緯にも依存すると考えるのは自然である。そして、その技術がどれだけ重要で画期的な内容を含んでいるか、利用するだけの価値があるかどうか大きく依存するものと考えられる。つまり、特許を出願はしてみたものの技術的にさほど高度な考案を含まない場合には審査請求を行っても拒絶される確率が高いし、コストをかけて権利を登録し維持する理由もあまりないため、利用するという意思決定には至らない。自社で独占するほどの技術ではないが競合他社に権利を抑えられるのを防ぐためには、出願して公知にしておくだけで通常は十分であろう。このような考えに基づき、個別特許の価値あるいは重要度に関係する指標を用いる。これらの指標としては、特許のスコープ（権利範囲の広さ）、早期審査請求の有無、引用や被引用の度合い（他の発明者や審査官により引用される回数は、それだけ先見的で技術的に重要な内容を含む特許である可能性が高い）、科学的知識への依拠など、複数のものが利用可能である（Suzuki, 2008, Zuniga et. al., 2009, など）。

一方、その技術を自社で利用するという意思決定がなされた後に、実際に利用が開始されるまでの期間は、アウトプット・ラグで述べたような研究プロジェクトの性格などに加えて、技術開発競争状態や製造能力、販売能力などの補完的資産をどの程度有しているかなどの因子に影響を受けることが考えられる。

すなわち、本分析では USE_{it} (i 分野・ t 期における R&D 成果を利用するかどうか | 0 = 利用しない、1 = 利用する) が R&D プロジェクトのサイズ ($RDsize_{it}$)、発明の経緯 (INV_{it})、特許スコープ ($SCOPE_{it}$)、早期審査の必要性 ($EARL_{it}$)、特許引用・被引用 (CIT_{it})、技術的な画期性 ($TVAL_{it}$)、科学的知識への依拠 (SL_{it}) の関数であり、利用の意思決定に至った R&D 成果については、 LRU_{it} (i 分野・ t 期における R&D 成果の出願から利用までのラグ) が R&D プロジェクトのサイズ ($RDsize_{it}$)、R&D プロジェクトの性格 ($RDchar_{it}$)、発明の経緯 (INV_{it})、技術的な画期性 ($TVAL_{it}$)、企業の規模 ($FIRM_{it}$)、競争状態 (COM_{it})、補完的資産 ($ASSET_{it}$) の関数であると仮定する。

$$USE_{it} = g(RDsize_{it}, INV_{it}, SCOPE_{it}, EARL_{it}, CIT_{it}, TVAL_{it})$$

$$LRU^*_{it} = h(RDsize_{it}, RDchar_{it}, INV_{it}, VAL_{it}, FIRM_{it}, COM_{it}, ASSET_{it})$$

$$LRU_{it} = \begin{cases} LRU^*_{it} & \text{if } USE_{it} = 1 \\ 0 & \text{if } USE_{it} = 0 \end{cases}$$

次節では、これらの仮説を検証するために、どのような指標を利用するのかを説明する。

4. 用いたデータと変数

本分析に用いたデータは、経済産業研究所が 2007 年に実施した発明者サーベイ（RIETI サーベイ）の個票データを基本としている。RIETI サーベイは日本に居住する発明者を対象として、その発明者がかかわったある特許を特定し、その発明とそれをもたらした研究開発プロジェクト、商業化の経緯などについて回答を求めたものである。全体で 5300 件に近い回答が得られている。RIETI サーベイのサンプル抽出方法や結果の概要については、長岡・塚田, 2007 を参照のこと。本分析では、全サンプル中から出願日もしくは優先権主張日が 2005 年以降のもの 5,250 件のみを対象とした。また RIETI サーベイの結果に加えて、特許データベースから得られる個別特許の統計的な変数を数種類用いて分析を行った。

(1) 研究開発のラグに関する変数

RIETI サーベイでは、特定の特許を特定の研究開発プロジェクトと関連付けるモデルを基本として採用している。このモデルでは、特定の研究開発プロジェクトを実施した成果として発明や特許が得られることを想定しており（図 1）、多くの計量経済モデルが採用しているような、人あるいは組織に持続的に蓄積される“知識ストック”が基になって成果が生まれるというモデルとはやや異なっていることに注意が必要である。このモデルに従えば、研究開発を開始して一定の時間が経過した時点で特許が出願される（特許出願は研究開発プロジェクトが終了しているかどうかにはよらない）。また、特許出願後一定の時間が経過した後にその技術の利用が開始される。サーベイでは、その研究開発プロジェクトの開始から特許出願までに要した時間（変数名：r_a_month）と、特許出願から利用開始までに要した時間（変数名：a_u_month）を発明者に 1 ヶ月単位で回答してもらっている。

データから見る限り、出願された特許のうち実際に利用されるものは半分程度しかないようである（残りの特許については「特許出願から利用開始まで」に要した時間に対する回答が得られていない）。

なお、“利用開始”の定義については質問票では特に詳しい定義は行わず回答者の解釈にゆだねられているが、当該箇所の質問文では以下のように表現されている：

問 6-3. 当該特許は、出願者によってその製品あるいは製造において利用されましたか

また、その直前の質問の回答用選択肢にも以下のような表現が用いられている：

問 6-2. 当該発明を特許化した理由

特許化された発明の自社製品・自社製造工程での排他的な利用

これらの質問票の表現から考えると、回答者が想定している“利用開始”とは、製品やサービスが実際に市場投入される前の段階で、試作品や製品の製造が開始された時点と認識されていた可能性が高いであろう。すなわち、製品やサービスの実際の売り上げや利益が発生するには、“利用開始”からさらにタイムラグが存在する可能性が高いが、そのラグの長さは本分析の対象外である。

【図 1 本分析が採用する研究開発と特許出願および商業利用の時系列モデル】

以下、それぞれのラグに関する変数の基本的性質について述べる。

① 研究開始から特許出願までのラグ

研究開始から特許出願までの期間 (r_a_month) の頻度分布を図 2 に示す。図 2 左はヒストグラムのビンの幅を 3 ヶ月とした場合であるが、明らかに 4 半期の周期性が認められる。図 2 右は 4 半期や半期の周期性を除くためにビン幅を 12 か月とした場合のヒストグラムである。分布形状としては、左側（小さい値）に偏った、指数分布ないしは幾何分布に近い。このような分布を示す変数の代表値としては、平均値より中央値のほうが適切であろう。表 1 を見ると、 r_a_month の中央値は 18 ヶ月であり、75 パーセンタイルが 24 ヶ月となっている。すなわち、このサンプルにおける代表的な特許は研究開始後 18 ヶ月で出願され、24 ヶ月以内に 77%の特許が出願されるということである。また出願までに 60 ヶ月以上かかっている特許は 3.5%程度しかない（参考表 1 参照）。

【図 2 研究開始から特許出願までの期間の分布（左：ビン幅 3 ヶ月、右：ビン幅 12 ヶ月）】

② 特許出願から利用開始まで

先にも述べたように、“利用開始”の回答が得られているのは 2,398 件の特許のみであり、全体の半分程度である。これらについて、特許出願から利用開始までの期間 (a_u_month) の頻度分布を図 3 に示す。図 2 と同様に左図はビンの幅を 3 ヶ月とした場合、右図はビン幅を 12 か月とした場合のヒストグラムである。この変数の分布形状としては、幾何分布よりも対数正規分布に近い。RIETI サーベイは 1995 年～2001 年に特許出願された特許を対象として、2007 年前半に実施されたものであるから、サーベイ実施までの出願後の経過時間は最短で 60 ヶ月、最長で 144 ヶ月である。すなわち、本変数は正確に言うと 60 ヶ月を超える期間のデータが一部切断（truncated）されている（図 4）。サーベイ実施時点では未だ利用が開始されておらず、サーベイ後に出願から 60 ヶ月以上経過して利用が開始される特許も多少はあるものと考えられるが、図 3 の分布形状が大きく変化するとは考えにくいであろう。

【図 3 特許出願から利用開始までの期間の分布（左：ビン幅 3 ヶ月、右：ビン幅 12 ヶ月）】

【図 4 出願年別の特許出願から利用までの期間の分布】

本変数の分布形状は、研究開始から特許出願までと同様に左に偏った分布であり、代表値としては中央値の 12 ヶ月を採用することが適当である。また 75 パーセンタイルは 24 ヶ月である。すなわち、このサンプルで利用に至る代表的な特許は出願後 12 ヶ月で利用が開始され、24 ヶ月以内に 75%の利用が開始される。

表 1 には参考のために、研究開始から利用開始まで (r_a_month と a_u_month の合計) の記述統計も示した。研究開始後 5 年～12 年で利用が開始される特許を通して見た場合、このサンプル中の代表

的な特許は研究開始から 33 ヶ月で利用が開始され、50 ヶ月以内には 75%の特許が利用される、ということである。

【表 1 研究開始から利用開始までの期間に関する 3 変数の記述統計】

なお、このようなラグ分布の性質から考えると、研究開発投資のストック化を有限ラグで計算する際には、従来用いられてきた、

$$K_t = (1 - \delta)K_{t-1} + R_{t-1}$$

に代えて、例えば

$$K_t = (1 - \delta)K_{t-1} + 0.21 \cdot R_t + 0.38 \cdot R_{t-1} + 0.21 \cdot R_{t-2} \quad (\text{特許出願をアウトプットとする場合：図 5})$$

もしくは、

$$K_t = (1 - \delta)K_{t-1} + 0.22 \cdot R_{t-1} + 0.22 \cdot R_{t-2} + 0.17 \cdot R_{t-3} \quad (\text{特許の利用をアウトプットとする場合})$$

などの定式化を用いることが適切であると考えられる（重み付けの配分根拠については参考表 1 を参照）。

【図 5 研究開発投資の支出期間と特許出願集計期間の関係】

(2) 研究開発のラグに影響を与える因子に関する変数

① 研究プロジェクトの規模に関する変数

ra_man_month6（研究開始から特許出願までの投入人・月）

ra_man（研究プロジェクトに関わった推定人数）：

RIETI サーベイでは、Q4.16 で研究開始から特許出願までの間に投入されたマンパワー（変数名：*ra_man_month6*）を調べている。図 6 左にこの設問に対する回答の分布を示す。この設問にはもともと 8 カテゴリーの選択肢が設定されていたが、本分析では 49 人・月以上の 3 カテゴリーを統合してトップコーディングし、6 カテゴリーとした。有効回答は 5,114 件であった。さらに、各カテゴリーに関して区間の中間値（それぞれ 2, 5, 9, 18, 36, 72 人・月）を代表値とみなし、先に説明した *r_a_month* で除すことにより、研究開発プロジェクトの推定参加人数（変数名：*ra_man*）を求めた。図 6 右図はその推定参加人数の分布である。1～2 名の少人数で実施される研究プロジェクトが圧倒的に多いことがわかる。

【図 6 研究開始から特許出願までに投入された人・月と推定投入人数の分布】

num_inv（発明者の数）： 本変数は、当該特許に記載された発明者の数である。値は 1 以上の正の整数であり、データは特許データベースから取得した（Goto and Motohashi, 2007）。

num_relp (関連特許の数) : 本変数は、当該研究開発プロジェクトから得られた全特許の数に関する、6 カテゴリーの順序変数である。データはサーベイの個票に由来する。

d_co_app (共同出願人の有無) : 本変数は、当該特許に共同出願人がいるかどうかを示すダミー変数である。データは特許データベースから取得した。

② 研究プロジェクトの性格に関する変数

d_basic (研究の範囲：基礎研究)、 **d_applied** (研究の範囲：応用研究) : これらの変数は、サーベイにおいて“当該発明につながった研究活動の範囲”を尋ねた設問の回答に基づくものである。選択肢は yes/no で複数選択が可能とされているので、基礎研究と同時に応用研究や開発が行われたケースも存在している。

d_poj_tech (研究目的：技術基盤強化) : 本変数は、サーベイにおいて“当該発明につながった研究の事業上の目的”を尋ねた設問の回答に基づくものである。この選択肢に yes と回答した発明者は、特定の既存事業や新規事業ではなく、当面の事業とは直結しない技術基盤の強化が目的であったと考えられる。

③ 発明の経緯に関する変数

d_main_prdc (研究プロジェクトの主目標)、 **d_sub_prdc** (予想可能な副産物)、 **d_no_rand** (研究によらない発明) : これらの変数は、サーベイにおいて当該発明が生まれた経緯を尋ねた設問への回答に基づくものである。元の選択肢のうち“発明のアイデアは通常業務から”と“研究開発費をかけずに直接特許化”を研究によらない発明としてまとめた。

d_prev_inv (先行発明への依拠) : 本変数は、サーベイにおいて当該特許が“先行特許を相当程度に基礎として生み出された”かどうかを尋ねた設問の回答に基づくものである。

④ 発明の性格に関する変数

d_product_only (発明内容：モノの発明)、 **d_process_only** (発明内容：方法の発明) : これらの変数は、サーベイにおいて当該発明が、モノの発明か、方法の発明か、それら両方かを尋ねた設問の回答に基づくものである。これらの選択肢は排他的に設定されている。

⑤ 技術的な画期性に関する変数

d_tri (三極出願かどうか)、 **d_std_vip** (標準・重要特許かどうか) : これらの変数は、サーベイの対象を抽出する際に基礎となった、特許の属性に基づくものである。三極出願をされるような特許や、標準・重要特許として選ばれるような特許には、技術的に高度なアイデアが含まれている可能性が高いと考えられる。

paper_3 (論文発表) : 本変数は、サーベイにおいて当該特許の元になった“研究開発の成果を科学技術論文としても公表したかどうか”を尋ねた設問の回答に基づくものである。本来は yes/no の日変数であるが、論文として公表される特許はかなり少ないので、他の設問の回答を組み合わせで no のカテゴリーの中でも“発明の技術水準が高くなかった”ものを別カテゴリーとし、3 カテゴリーの順序変数とした。

⑥ 企業変数

firm_size (企業規模) : 本変数は、サーベイにおいて発明当時に所属していた企業の規模を尋ねた設問への回答に基づく、4 カテゴリーの順序変数である。

⑦ 技術分野固定効果

d_chem (米国分類：化学分野)、 *d_comp* (米国分類：情報通信分野)、 *d_drug* (米国分類：医薬・医療分野)、 *d_elec* (米国分類：電気・電子分野)、 *d_mech* (米国分類：機械分野)：これらの変数は、当該特許に付与された技術分類(米国分類に基づく)に基づく、排他的なダミー変数である。基準分野は“その他分野(ソフトウェア技術を含む)”とした。

⑧ 年度固定効果

d_1996 (1996 年出願・優先権主張) など： これらの変数は、当該特許の出願年もしくは優先権主張年の早いほうを示す、排他的なダミー変数である。基準年は 1995 年とした。

(3) 特許出願から商業的利用までの期間に関する説明変数

特許出願後に利用の意思決定が行われる際の判断基準と、実際に利用が開始されるまでの期間に影響を与える因子の変数としては、上で述べた研究開発から特許出願までの期間に関する説明変数以外にも、発明の価値や当該技術分野における競争状況、特許と組み合わせられる補完的な企業戦略などに関する変数を用いた。また、研究開発のアウトプット・ラグも説明変数に加えている。

① 特許のスコープに関する変数

num_ipc (IPC サブクラス数)： 本変数は、当該特許に付与された IPC 技術分類のサブクラス(4 桁)レベルの数を示す変数である。発明が依拠するあるいは応用可能な技術分野のスコープを示すと解釈できる。データは特許データベースから取得した。

famsize (ファミリーサイズ)： 本変数は、当該特許が属するパテントファミリーの大きさ(外国出願の国の数)を示す変数である。三極出願はファミリーサイズが 3 で日欧米で出願・登録されたものだけを指すが、本変数では自国に加えて特定の外国 1 国のみの出願や、4 カ国以上への出願なども識別される。データは特許データベースから取得した (PATSTAT, 2010)。

ln_claim (請求項の数)： 本変数は、当該特許の出願時の請求項の数を示す変数である。データは特許データベースから取得した。

② 特許の早期審査の必要性に関する変数

d_earl_ex (早期審査請求の有無)： 本変数は、当該特許が出願された後 1 年以内(出願と同時に含む)に審査請求が行われたかどうかを示すダミー変数である。データは特許データベースから取得した。

③ 特許の引用・被引用に関する変数

inv_fc5_or (発明者 5 年前方引用)： 本変数は、当該特許が出願されてから 5 年以内に、後発特許の発明者によって引用された回数を示す変数である。データは特許データベースから取得した (Suzuki, et.al., 2008)。

exm_fc5_or (審査官 5 年前方引用)： 本変数は、当該特許が出願されてから 5 年以内に、審査官によって引用された回数を示す変数である。データは特許データベースから取得した。

inv_bc (発明者後方引用数)： 本変数は、当該特許中に引用されている先行特許文献の数を示す変数である。データは特許データベースから取得した。

④ 科学的知識への依拠に関する変数

num_sl (サイエンスリンケージ数) : 本変数は、当該特許中に引用されている先行科学技術文献の数を示す変数である。データは特許データベースから取得した。

⑤ 技術分野の競争状態に関する変数

itc_total (技術分類サイズ) : 本変数は、当該特許が属する技術分類中分類 (33 技術分野) への、日本国内における一年間の総出願数を示す変数である。データは特許データベースから取得した。

itc_conc (技術分類集中度) : 本変数は、当該特許が属する技術分類中分類 (33 技術分野) に日本国内で出願された特許のうち、上位 3 位までの出願人による出願の比率 (上位 3 社集中度) を示す変数である。データは特許データベースから取得した。

⑥ 技術の商業化にとって重要な企業戦略に関する変数

asset_tech (戦略 : 技術開発先行)、 **asset_lead** (戦略 : 市場投入先行)、 **asset_manu** (戦略 : 製造能力)、 **asset_sales** (戦略 : 販売能力)、 **asset_enfor** (戦略 : 特許権エンフォース)、 **asset_know** (戦略 : 企業秘密)、 **asset_cmplx** (戦略 : 複雑性)、 **asset_coop** (戦略 : 外部企業協力) : これらの変数は、サーベイにおいて“当該発明の商業的成功のための企業戦略”として、それぞれがどの程度重要であるかを尋ねた設問の回答に基づくものである。回答はそれぞれの戦略の重要度を独立に 5 段階で聞いており、順位付け等を行っていない。

参考表 2 に、これら変数の定義をまとめる。また参考表 3 は、説明変数に関する記述統計である。

5. 推計結果と考察

(1) 研究開発から特許出願までの期間

研究開発の開始から特許出願まで(r_a_month)に関する説明変数の係数を GLM によって推計した。被説明変数の分布として仮定したのは幾何分布(=負の二項分布；パラメータ 1)で、リンク関数に Log を採用した。幾何分布の統計的な解釈としては、一般的に「生起確率 P の独立ベルヌーイ試行を繰り返し、初めて成功するまでの試行回数」とされており、この場合の“試行回数”を“研究開発開始後の月数”と置き換えると理論的にも矛盾無く理解することができる。表 2 に推計結果を示す。

【表 2 研究開発の開始から特許出願までの期間に関する GLM 推計結果】

研究プロジェクトの規模は num_inv や num_relpat の係数がプラスで有意となっており、やはり大規模な研究開発プロジェクトでは、研究開始から特許出願までの時間が長くなることを示しているものと考えられる。一方、ra_man の係数はマイナスで有意となった。表 3 に、これら 3 変数の相関マトリクスを示す。相関はさほど強くないが、単相関の係数はいずれもプラスであり、いずれの指標も理論的に研究プロジェクトの規模と正の相関を持つものと考えられる。先に述べたように ra_man は順序カテゴリー変数とスケール変数から計算された半定量的な変数であり、その解釈には十分な注意が必要であるが、同じようなプロジェクトで比較した場合、多人数で取り組む研究開発プロジェクトでは、少人数のプロジェクトよりも短期間で成果が得られる（発明にかかる時間と労力は代替的な関係にある）ことを示唆しているものと考えられる。

【表 3 研究プロジェクトの規模に関する説明変数の相関マトリクス】

共同出願人の存在 (d_co_app) は、有意性は高くないものの予想通り出願までの時間を長引かせる方向に影響を与えるようである。

研究開発プロジェクトが基礎的な内容を含むかどうか (d_basic、d_applied) や、既存事業と関係ない技術基盤の強化 (d_poj_tech) を狙ったものかどうかは、アウトプット・ラグには影響を与えないことが示された。これはやや意外な結果であるが、見方を変えるなら企業の研究者にとって特許出願そのものが研究開発プロジェクトの目的であり、事業化とは関係ない基礎的な研究であつてもなるべく早く特許を出願することが求められるということを反映しているのかもしれない。

上述の研究プロジェクトの性格と同様に、発明の経緯が当然予期されたものであったかセレンディピティによるものかは、アウトプット・ラグに影響を与えないようである (d_main_prdc、d_sub_prdc)。ただし、研究開発に拠らない発明の場合 (d_no_rand) には当然のことながら出願までの時間は短い。また、先行発明に依拠する場合 (d_prev_inv) に出願までに長い時間がかかっているのは予想通りである。これは、当該研究開発プロジェクトが、先行発明に関係するプロジェクトからの継続であるとみなされた可能性が高いだろう。

発明の性格が、新プロセスに関係せず新プロダクトのみに関わる（`d_product_only`）ようなものであった場合には、出願までの時間がやや短くなる傾向が見られる。これは、プロセスに関する発明は特許で保護せずに、ノウハウで保護することも可能であるということが反映されていることが考えられる。

三極特許（`d_tri`）は、しばしば重要な特許の代名詞として利用されるが、分析の結果それらの特許では出願までの時間がやや長くなることが示された。また、論文発表（`paper_3`）の係数がプラスで有意性が大きい、民間企業においては論文発表を優先して特許出願を遅らせることは考えにくいので、これは発明の内容が科学技術論文として発表されるような成果は、研究開始から出願までに長い時間を要する規模の大きなプロジェクトから得られることが多いものと解釈するべきであろう。

企業規模の変数（`firm_size`）の係数はマイナスで有意であり、基準となっている大企業に対して中小企業では出願までの時間が短いことを示している。これは、知財を扱う専門部門の支援がないような中小企業でも、外部の特許事務所等の支援を得れば、少なくとも特許出願を迅速に行うことができるということを意味しているのかもしれない。

技術分野に特有の固定効果については、化学分野と医薬品分野の係数がプラスで有意であった。これは、特に医薬品分野では研究開発に長い時間がかかるという、一般的な認識と整合的である。

一方、年度の固定効果は、いずれの出願年についても有意ではなかった。すなわち、1995年～2001年について見る限り、研究開発のアウトプット・ラグが時代とともに短縮するような傾向も、逆に長くなるような傾向も見られない。一般的には、近年、研究開発に早急な成果が求められるようになり、研究開始の期間が短縮しているのではないかという見方があるが、そのような見方を支持する証拠は本分析では得られなかった。

(2) 特許出願から利用開始までの期間

サンプルセレクション・バイアス

本分析では、特許出願から利用開始までの期間（`a_u_month`）に関するデータを打ち切りデータと見て、説明変数に含まれないサンプルの効果を取り入れたヘックマンの2段階推定を行った。すなわち、第1段階は“特許出願から利用開始まで”の期間を回答しているかどうか、すなわちその特許が利用されているかどうかの二値変数（`d_au_month` | 0=利用していない、1=利用されている）を被説明変数とするプロビット推定である。説明変数には第2段階で使用する説明変数を考慮し、第2段階のラグ関数に大きな影響を与えない、発明の経緯や特許の技術的価値に関する指標を中心に使用した。そして、それにより推定したパラメータおよび逆ミルズ比を用いて、第2段階のOLS推定を行った。第2段階の被説明変数は、対数変換した出願から利用までのラグ（`log_au_month`）である。表4に、ヘックマンの2段階推定の結果と、比較のために通常のOLSの結果を示す。

【表4 特許出願から利用開始までの期間に関する推計結果】

表4からわかるように、逆ミルズ比は非常に小さく統計的に有意ではない。またヘックマンの2段階推定の結果と通常のOLSの結果が、係数の符号や値、有意性などほとんど差がなく、サンプルセレクションによるバイアスはほとんどないという結論が得られた。すなわち、“利用されるかどうか”と“出

願から利用までの期間”に寄与する因子は独立性が高く、「すぐ利用できそうなものを選んで利用している」というような傾向は考えにくいということである。

当該特許を利用するかどうか

研究開発のアウトプット・ラグである、研究開始から特許出願まで (\log_ra_month) の係数の符号はマイナスで有意である。すなわち、アウトプット・ラグの長い特許は、利用されにくいことを示している。また同様に、当該発明が研究開発プロジェクトの予想可能な副産物であったり、研究開発を経ない発明であったりするものは利用されにくい。これらをあわせて考えると、最も利用される可能性の高い特許は、研究開発プロジェクトの主たる目標もしくはセレンディピティによる発明で、プロジェクト開始後早期に出願されたものということになる。民間企業における研究開発では、スピードが重要視されることを示しているものと考えられる。

特許のスコープに関する変数は、IPC の数 (num_ipc)、請求項の数 (ln_claim) とともに係数がマイナスで有意であった。出願時の IPC 数や請求項数が多い特許は、逆に言うところと応用先があまりはっきりしていない基礎的な発明であり、利用されにくいということを示しているのかもしれない。一方、ファミリーサイズ ($famsize$) と三極特許 (d_tri) の係数はプラスで有意性も高い。外国出願にはかなりのコストがかかることから、それだけのコストをかけてまで出願される特許はビジネス上の価値が高く、利用される可能性が高いと考えてよいだろう。

特許引用に関係する変数は、発明者被引用 (inv_fc5_or) と審査官被引用 (exm_fc5_or) の係数が共にプラスで有意であった。特に、後続の発明者によって多く引用されるような特許は、商業的にも利用される可能性が高いことを示している。ただし、被引用の集計はある程度の時間が経過してからでないとできないため、これらの指標を予測的に用いることには限界がある。

研究開発の成果を論文として発表したかどうか ($paper_3$) の係数はプラスで、絶対値が最も大きく、有意性も高い。科学的に高度な知見を含む発明は、利用される可能性が高いことを示している。一方、特許中に科学技術論文が引用されていることを示すサイエンス・リンケージ (num_sl) の係数はマイナスで、絶対値が大きく有意性も高い。科学的知見への依拠が強い発明は基礎的で、商業的利用にはあまり向かないことを示しているものと考えられる。ただし、表 5 に示すように、論文発表とサイエンス・リンケージは独立ではなく、サイエンス・リンケージのある特許は論文発表される確率が高い。すなわち、両指標とも研究開発プロジェクトと科学的知識との関係を示しているのは確かだと考えられるが、サイエンス・リンケージはインプットとして利用する知識の性質を表し、発明者による論文発表はアウトプットとして生み出された知識の性質を表しているという点で、それらの持つ意味は異なっている。筆者は、これらの指標は後述する組織の吸収能力 ($absorptive\ capacity$) や産学連携との関係が深いのではないかと考えているが、詳しい分析は残された課題である。

【表 5 論文発表とサイエンス・リンケージの関係】

技術分野に特有の固定効果については、特に医薬品分野とエレクトロニクス分野の係数がマイナスで有意性が高い。医薬品分野とエレクトロニクス分野では、特許は出願したもののその後の開発で放棄さ

れたり、あるいはそもそも純粋な防衛目的で出願されるものが多かったりするという状況を反映しているものと考えられる。

出願された特許が利用されるかどうかについて、年度の固定効果はいずれの出願年についても有意ではなかった。すなわち、出願はしたが利用されない特許は、常に一定の割合で存在しており、その比率に顕著な変化は見られないということである。

出願から利用開始までのラグ

研究開発のアウトプット・ラグ (\log_ra_month) の係数は、プラスで有意であった。また、関連特許の数 (num_relpat) の係数もプラスで有意である。研究開始から特許出願までに長い時間がかかり、関連特許の数も多いような規模の大きな研究開発プロジェクトの成果は、利用開始までにも長い時間がかかることを示している。一方、共同出願人がある特許 (d_co_app) は、利用開始までの時間がやや短くなることが示唆された。

研究プロジェクトの性格として基礎的な内容を含む (d_basic) 場合には、利用までの時間が短くなるという、当初の想定とは逆の結果が得られた。サーベイでは基礎／応用／開発的な要素を含むかどうかは重複回答を可能としているため d_basic や $d_applied$ は排他的なダミー変数ではない。すなわちこの変数が 0 の値をとるのは、研究開発プロジェクトに基礎的な内容が含まれない、応用もしくは開発に特化したプロジェクトの場合が多いものと考えられる (図 7)。このようなプロジェクトの例としては、外部で生まれた基本的なアイデアを導入し製品化するようなものが典型として考えられよう。すなわち、外部の基礎研究で開発された成果を導入するような場合には、その技術を理解し、自社のものとして利用することが可能となるまでには時間がかかるということを示しているのかもしれない。これは組織の吸収能力にも関係する知見であり、今後の詳しい検証が望まれる。

【図 7 研究プロジェクトの性格に関するダミー変数間の包含関係】

発明の経緯 (予想された結果かどうか) や、発明の性格 (モノの発明か方法の発明か) は、利用までのラグには明確な影響を及ぼさないようである。

発明内容の技術的な価値と関係の深い 3 種類の変数 (d_tri 、 d_std_vip 、 $paper_3$) の係数は、いずれもプラスで有意であった。すなわち、技術的に画期的な成果は、利用されるまでに長い時間を要することを示しているものと考えられる。

技術分野の競争状態を示す変数 (itc_total 、 itc_conc) の係数は、予想に反して有意とはならなかった。利用までの時間に影響を与えるのは、技術的な競争環境ではなくマーケットシェア等の市場競争環境であるということかもしれない。

当該発明の商業的成功のために重要な企業戦略としては、“技術的な先行優位 ($asset_tech$)” と “優れた製造能力 ($asset_manu$)”、“外部との連携 ($asset_coop$)” の係数がプラスで有意であった。一方、“市場における先行優位 ($asset_lead$)” の係数はマイナスで有意であった。すなわち、市場で先んじることが重要な成果については実際に迅速に利用されていることを示しており、予想通りの結果である。

企業規模 ($firm_size$) の係数はマイナスで有意である。これは大企業に比べて中小企業では意思決定が早く、新しい技術の導入も素早く行えるということを示しているものと考えられる。

技術分野に特有の固定効果は、化学分野（`d_chem`）および医薬品分野（`d_drug`）の係数がプラスで有意である。これらの分野では予想通り、特許出願から利用開始までにもさらに長い時間が必要であることを示している。

年度の固定効果は、1998 年以降の係数がマイナスで有意であった。これを単純に解釈すると“特許出願から利用開始までの期間が近年短縮している”となり、「研究開発の加速現象」を支持する結果とも考えられる。しかし、変数 `a_u_month` の項で述べたように、利用開始までの期間はデータの制約上、部分的に切断されている。すなわち、2001 年出願の特許では 60 ヶ月以上の部分が、2000 年出願の特許では 72 ヶ月以上の部分が、1999 年出願の特許では 84 ヶ月以上の部分が、1998 年出願の特許では 96 ヶ月以上の部分が本データでは測定されておらず、係数に下方バイアスを与えている可能性がある。それではデータ切断の影響が実際に係数に現れているのだろうか？出願年による `a_u_month` の分布の違いは、図 4 に示したように、主に出願後間もない 12 ヶ月～24 ヶ月の部分の分布に現れているようである。また年度固定効果の係数の絶対値を見ると、年度とともにスムーズに増大するというよりも 1998 年以降が不連続に増大している²。これらを考慮すると、年度による違いは徐々に進行する「研究開発の加速」や切断によるバイアスというよりも、何らかの制度変更による構造変化を疑わせる結果となっている。この頃の大きな制度変更としては、1997 年 10 月より実施された審査請求期間の 7 年から 3 年への短縮があげられる。特許庁によるとこの制度変更により審査請求率が 1998 年出願の特許以降、顕著に増加したことが報告されている（特許行政年次報告書, 2010 年）ので、これが発明者の認識する“利用開始”時期に影響を与えた可能性があるだろう。ただし、分析対象期間が 7 年間と短い上に、データ切断のバイアスの可能性も残ることから、実際に利用開始までの期間が短縮しているのかどうか、本分析結果をもって結論付けることは危険であろう。

² データ切断の影響を除外するため、特許出願から利用開始まで（`a_u_month`）が 60 ヶ月を超えるケース（2,148 のうち 67 ケースが相当）を全て除いた推計も行った。その結果、それぞれの年度固定効果の係数の有意性は低下したものの符号は同一で、1998 年以降の係数の絶対値が不連続に大きく増大するという傾向は不変であった。

6. まとめ

研究開発ラグの分布について

研究開発活動に対するインプットからアウトプット、アウトカムまでにいたるタイムラグは、いくつかの性格の異なる期間が合成された結果であると考えられる。本分析では、企業に所属する発明者というおそらくはある程度の技術的知識ストックを持っているだろう研究者が、新たな研究開発プロジェクトを開始してから特許を出願するまでのラグと、特許を出願した後に利用開始に至るまでのラグを調べた。研究開始から特許出願までのラグは、コイック・ラグが想定する幾何分布でさほど矛盾がなく、中央値は 18 ヶ月であり 24 ヶ月までには 77%のサンプルが収まるという結論を得た。また、特許出願後にその技術を利用するかどうかの意思決定が行われ、利用の意思決定が行われたものに関しては出願後 12 ヶ月（中央値）で利用が開始される。特許出願後利用開始までのラグは、さらに利用の意思決定までの期間と利用決定後実際に利用が開始されるまでの期間の合成期間であると考えられるが、本分析ではデータの制約上それらを区別して論じることができない。特許出願後利用開始までのラグ分布は、対数正規分布に近い形状を有しており、24 ヶ月までに 75%のサンプルが収まる。上記のラグを通して見た場合には、中央値は 33 ヶ月、75 パーセンタイルは 50 ヶ月である。

特許出願後利用開始までの期間

出願した特許を利用するかどうかの意思決定に寄与する因子と、出願後利用開始までの期間に影響を与える因子は独立性が高く、利用の意思決定は主として発明内容が技術的に高い価値を持つかどうかに依存する。一方、利用開始までのラグは、主として研究プロジェクトの規模や企業戦略により大きな影響を受けることがわかった。

ラグに影響を与える因子について

その他、ラグに影響を与える主な因子としては、まず技術分野の影響がある。化学分野の発明は、研究開始から出願までの期間が約 12%（2 ヶ月）長く、医薬品分野では約 24%（4 ヶ月）長い。また出願から利用開始までの期間も、化学分野では約 26%（3 ヶ月）長く、医薬品分野では約 20%（2 ヶ月）長い。これらの分野では一般的に研究開発に長期間が必要であると認識されている。確かに、本分析結果はそのような認識と合致するが、12%から 26%という違いは、見方によってはそれほど大きくないともいえる。これは、例えば医薬品の研究開発を考えた場合、時間がかかるのは主に前臨床試験や臨床治験、毒性試験などのフェーズであり、これらのフェーズでは既に新技術は試験製品の製造のために“利用が開始”されているためであると考えられる。すなわち、利用開始後に実際に製品が市場に導入され、売上や利益が計上されるようになるまでには、さらに別種のタイムラグを考える必要があるだろう。

中小企業は大企業に比べて、研究開始から特許出願までの期間も、特許出願から利用開始までの期間も短いことが示された。

研究開発期間の短縮や加速化などの現象が、実際に存在するかどうかを明確に示すことはデータの制約もありできなかった。ただし、研究開始から特許出願までのラグについては、調べた期間で変化は見られず、研究開発の加速化はあったとしてもそれほど顕著なものではないと考えられる。また、出願か

ら利用開始までのラグは、1998年以降不連続に短縮しており、1997年に導入された審査請求期間の短縮が影響を与えている可能性がある。

基礎的な要素を含む研究開発のラグが短いことや、科学とのリンケージの強いプロジェクトの成果が利用され難いことなど、やや事前の予想とは異なる結果が得られた。これらの点については、産学連携との関係や組織の技術的知識吸収能力との関係をあわせて、今後の検討が待たれる。

参考文献

- Corrado, C. A., Hulten, C. R. & Sichel, D. E. , 2006, Intangible Capital and Economic Growth. NBER Working Paper No.11948.
- Goto and Motohashi, 2007, Construction of a Japanese Patent Database and a first look at Japanese patenting activities, Research Policy, Volume 36, Issue 9.
- Griliches, Z. ed. (1984), R&D, Patents and Productivity, Chicago, The University of Chicago Press.
- National Academies, 2007, Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Future, Committee on Prospering in the Global Economy of the 21st Century, The National Academies Press, Washington, D.C.
- Mowery, David C. (1994), Science and Technology Policy in Independent Economy, Kluwer Academic Publishers.
- Suzuki, 2008 年, Structural Modeling of the Value of Patent, RIETI Discussion Paper Series 08-J-039, 経済産業研究所.
- Suzuki, et.al., 2008, Patent Citations in Japan – Database Construction for Inventor Citations and Examiner Citations, Conference on Patent Statistics for Decision Makers, OECD/EPO, Sept. 3, 2008, Vienna.
- PATSTAT, 2010, EPO Worldwide Patent Statistical Database, April 2010 version, European Patent Office.
- Zuniga, P. et. al., 2009, OECD Patent Statistics Manual, OECD.
- 科学技術政策研究所 第 1 研究グループ, 1999 年, 研究開発関連政策が及ぼす経済効果の定量的評価手法に関する調査(中間報告), NISTEP REPORT No. 64, 科学技術庁.
- 科学技術政策研究所 第 2 研究グループ, 2010 年, 民間企業の研究活動に関する調査報告, NISTEP REPORT No.143, 文部科学省.
- 元橋一之、2009 年、「日本企業の研究開発資産の蓄積とパフォーマンスに関する実証分析」、深尾京司編『マクロ経済と産業構造』、慶応大学出版会、pp 251-288.
- 後藤晃・本城昇・鈴木和志・滝野沢守、1986 年、「研究開発と技術進歩の経済分析」、経済分析, 第 103 号、昭和 61 年 9 月 経済企画庁経済研究所.
- 特許行政年次報告書, 2010 年, 産業財産権の現状と課題～125 周年を迎えた産業財産権制度～, 特許庁.
- 長岡貞男・塚田尚稔, 2007 年, 発明者から見た日本のイノベーション過程: RIETI 発明者サーベイの結果概要 1, RIETI Discussion Paper Series 07-J-046, 経済産業研究所.
- 横山重宏・吉本陽子, 2005 年, 研究開発促進税制の経済波及効果に係る調査報告書, 平成 16 年度産業技術調査, 株式会社 UFJ 総合研究所.

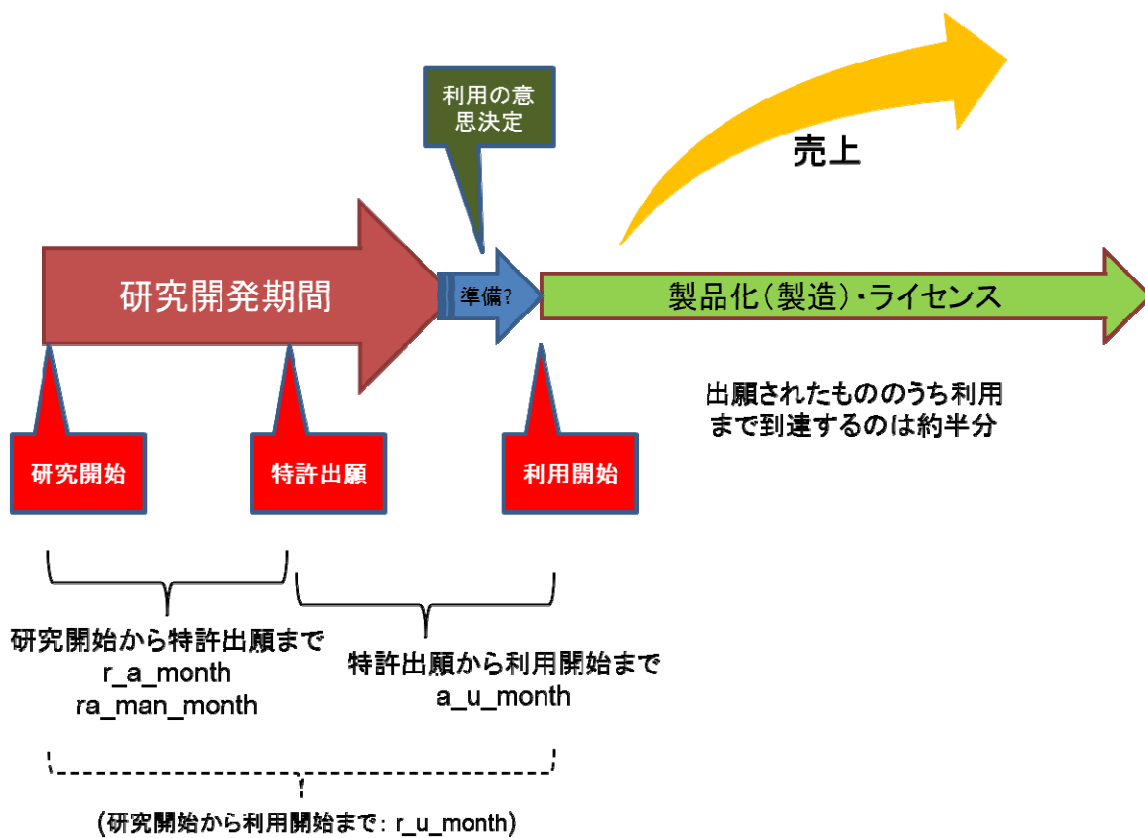


図1 本分析が採用する研究開発と特許出願および商業利用の時系列モデル

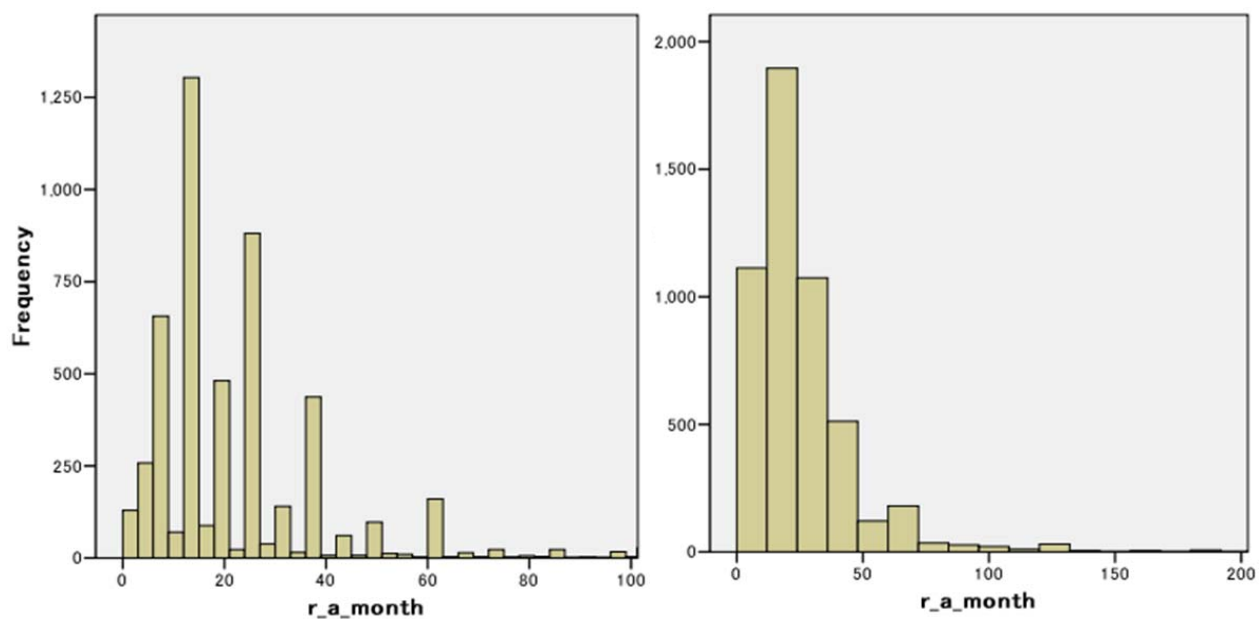


図2 研究開始から特許出願までの期間の分布（左：ビン幅3ヶ月、右：ビン幅12ヶ月）

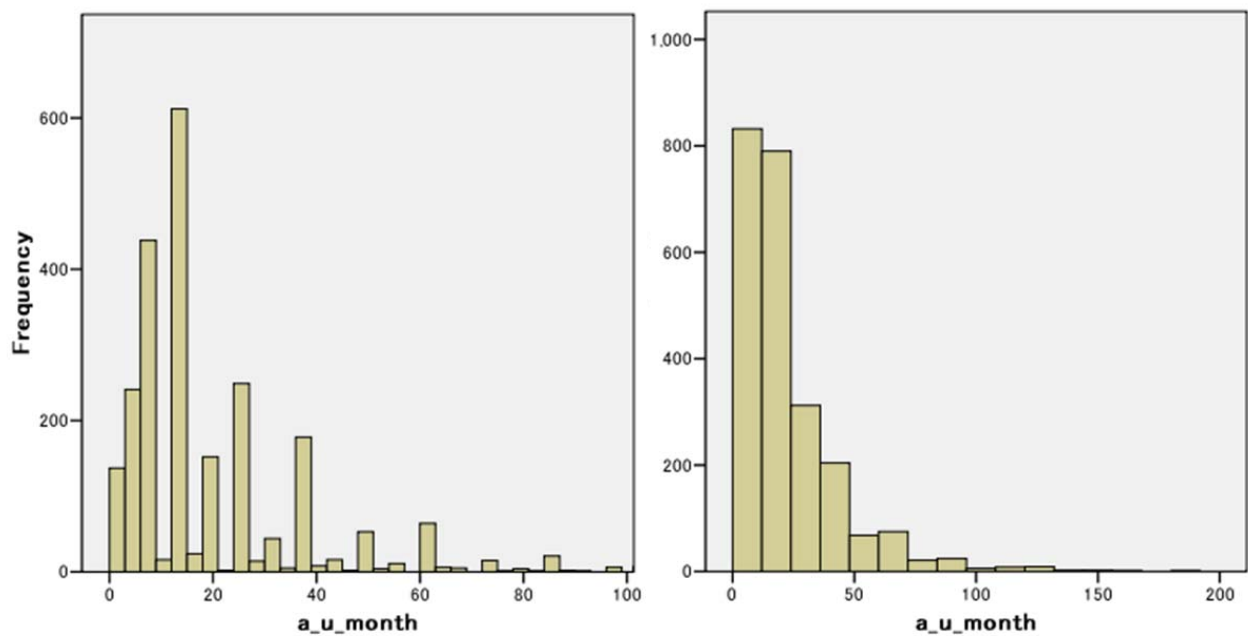


図3 特許出願から利用開始までの期間の分布（左：ビン幅3ヶ月、右：ビン幅12ヶ月）

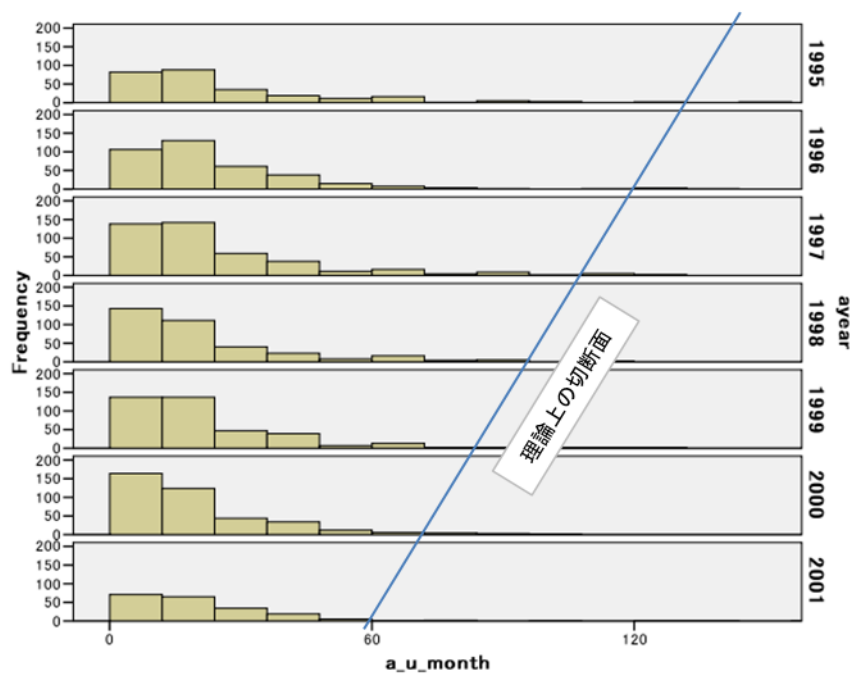


図 4 出願年別の特許出願から利用開始までの期間の分布

発明者サーベイが実施された時点までの期間を“理論上の切断面”として図中に示した。

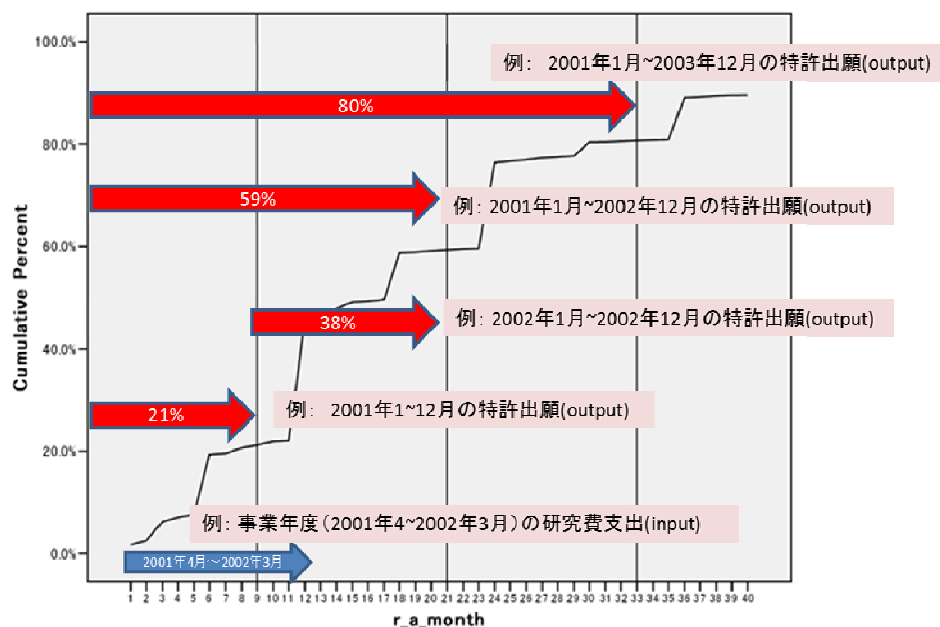


図 5 研究開発投資の支出期間と特許出願集計期間の関係

2001 年度の研究開発支出 input に対してカレンダー年で特許出願 output を集計する場合を例として示した。

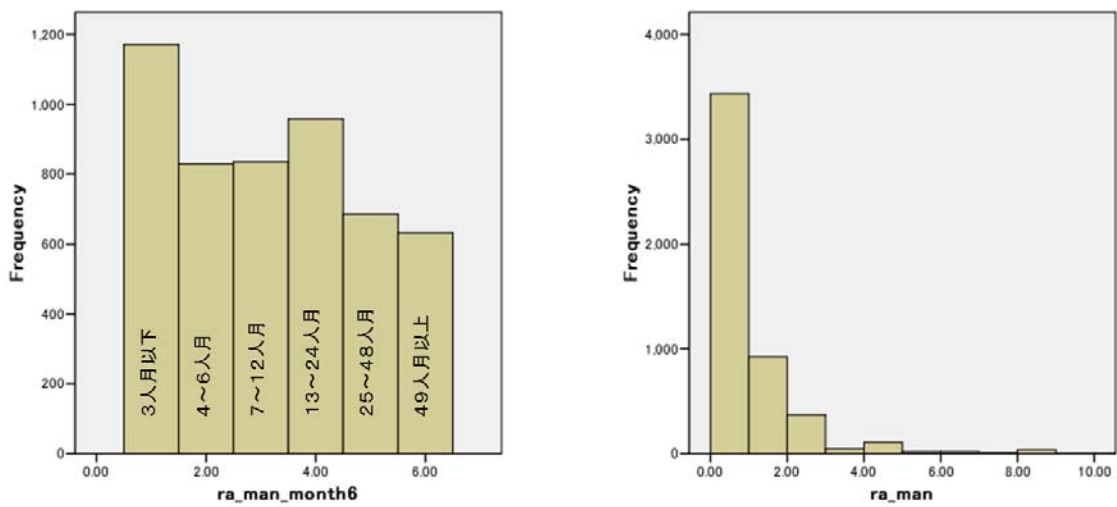


図 6 研究開始から特許出願までに投入された人・月と推定投入人数の分布

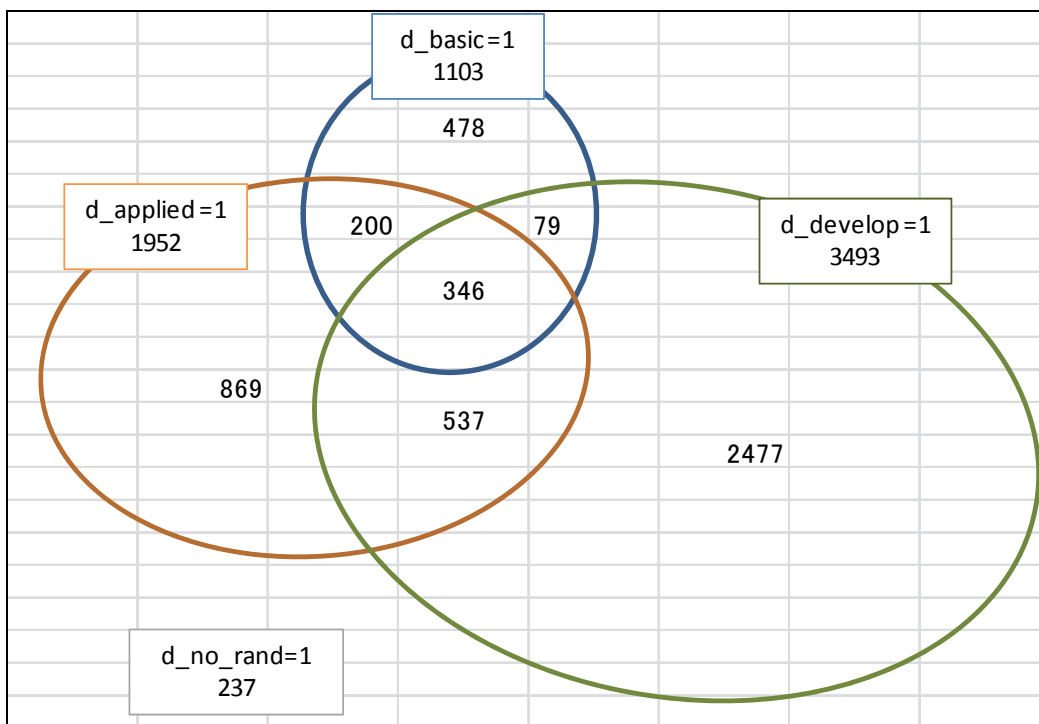


図 7 研究プロジェクトの性格に関するダミー変数間の包含関係（推計では $d_develop$ は用いていない）

表 1 研究開始化から利用開始までの期間に関する 3 変数の記述統計

	r_a_month 研究開始から出願	a_u_month 出願から利用開始	r_u_month 研究開始から利用開始
サンプル数	5051	2398	2357
平均値	22.49	19.32	41.63
中央値	18	12	33
最頻値	12	12	24
標準偏差	25.056	21.447	37.011
最少	1	1	2
最大	480	288	642
パーセンタイル			
5%	3	2	9
10%	6	3	12
25%	12	6	18
50%	18	12	33
75%	24	24	50
90%	42	42	80.2
95%	60	60	102

表 2 研究開発の開始から特許出願までの期間に関する GLM 推計結果

	GLM estimates			
	Probability Distribution	Negative binomial (1)		
	Link Function	Log		
	Number of obs	=	4,630	
	Likelihood Ratio Chi-Square	=	508.56	
	Prob > chi2	=	0	
	被説明変数	r_a_month		
	変数	係数	標準誤差	
研究プロジェクトの規模				
	ra_man	-0.111	0.009	***
	num_inv	0.036	0.010	***
	num_relpat	0.149	0.015	***
	d_co_app	0.082	0.044	*
研究プロジェクトの性格				
	d_basic	0.060	0.038	
	d_applied	0.010	0.032	
	d_poj_tech	0.017	0.056	
発明の経緯				
	d_main_prdc	0.046	0.045	
	d_sub_prdc	-0.065	0.049	
	d_no_rand	-0.330	0.057	***
	d_prev_inv	0.090	0.031	***
発明の性格				
	d_product_only	-0.075	0.040	*
	d_process_only	-0.043	0.042	
技術的な画期性				
	d_tri	0.078	0.037	**
	d_std_vip	0.025	0.121	
	paper_3=1	0.060	0.072	
	paper_3=2	0.298	0.081	***
企業変数(企業規模)				
	firm_size=2	-0.243	0.101	**
	firm_size=3	-0.124	0.092	
	firm_size=4	-0.259	0.069	***
技術分野固定効果				
	d_chem	0.124	0.048	**
	d_comp	-0.051	0.052	
	d_drug	0.244	0.071	***
	d_elec	0.035	0.049	
	d_mech	0.018	0.047	
年度固定効果				
	d_1996	-0.050	0.060	
	d_1997	-0.085	0.059	
	d_1998	-0.094	0.061	
	d_1999	-0.048	0.060	
	d_2000	-0.053	0.059	
	d_2001	-0.036	0.070	
定数項				
		2.791	0.120	***
***: 1% sig., **: 5% sig., *: 10% sig.				

表 3 研究プロジェクトの規模に関する説明変数の相関マトリクス

	ra_man	num_inv	num_relpat
ra_man	1		
num_inv	0.082	1	
num_relpat	0.147	0.078	1

表 4 特許出願から利用開始までの期間に関する推計結果

Heckman selection model -- two-step estimates (regression model with sample selection)					OLS estimates		
Number of obs		=	4,727		Number of obs	=	2,148
Censored obs		=	2,599		F(37, 2110)	=	12.09
Uncensored obs		=	2,128		Prob > F	=	0
Wald chi2(37)		=	452.8		R-squared	=	0.175
Prob > chi2		=	0		Adj R-squared	=	0.160
被説明変数	d_au_month (PROBIT)		log_au_month		log_au_month		
説明変数	係数	標準誤差	係数	標準誤差	係数	標準誤差	
研究プロジェクトの規模							
log_ra_month	-0.049	0.023 **	0.254	0.026 ***	0.254	0.026 ***	
num_relpat			0.117	0.020 ***	0.111	0.020 ***	
d_co_app			-0.095	0.056 *	-0.091	0.056	
研究プロジェクトの性格							
d_basic			-0.109	0.050 **	-0.107	0.050 **	
d_applied			0.011	0.041	0.015	0.042	
d_poj_tech			0.260	0.115 **	0.263	0.116 **	
発明の経緯							
d_main_prdc	0.014	0.055	-0.024	0.056	-0.027	0.057	
d_sub_prdc	-0.210	0.061 ***	-0.046	0.068	-0.038	0.065	
d_no_rand	-0.334	0.070 ***	-0.074	0.083	-0.058	0.077	
d_prev_inv	-0.072	0.038 *	0.061	0.042	0.074	0.041 *	
発明の性格							
d_product_only			0.010	0.053	0.011	0.053	
d_process_only			-0.064	0.056	-0.064	0.056	
特許スコープ							
num_ipc	-0.046	0.007 ***					
ln_claim	-0.063	0.025 **					
famsize	0.053	0.010 ***					
特許早期審査							
d_earl_ex	0.094	0.058					
特許引用・被引用							
inv_fc5_or	0.094	0.026 ***					
exm_fc5_or	0.065	0.033 **					
inv_bc	-0.002	0.006					
技術的な画期性							
d_tri	0.184	0.060 ***	0.144	0.063 **	0.102	0.052 **	
d_std_vip	0.054	0.152	0.283	0.162 *	0.230	0.156	
paper_3	0.365	0.046 ***	0.266	0.062 ***	0.232	0.056 ***	
科学的知識への依拠							
num_sl	-0.316	0.041 ***					
技術分野の競争状態							
itc_total			0.000	0.000	0.000	0.000	
itc_conc			-0.002	0.430	0.036	0.432	
補完的な資産							
asset_tech			0.129	0.030 ***	0.131	0.030 ***	
asset_lead			-0.109	0.031 ***	-0.114	0.031 ***	
asset_manu			0.070	0.030 **	0.072	0.030 **	
asset_sales			-0.034	0.027	-0.037	0.027	
asset_enfor			-0.007	0.029	-0.007	0.029	
asset_know			-0.018	0.027	-0.018	0.027	
asset_cmplx			0.003	0.027	0.000	0.027	
asset_coop			0.043	0.024 *	0.042	0.024 *	
企業変数(企業規模)							
firm_size			-0.054	0.025 **	-0.056	0.025 **	
技術分野固定効果							
d_chem	-0.113	0.062 *	0.260	0.069 ***	0.292	0.066 ***	
d_comp	-0.109	0.064 *	0.025	0.075	0.033	0.074	
d_drug	-0.395	0.092 ***	0.201	0.113 *	0.262	0.102 **	
d_elec	-0.131	0.060 **	0.056	0.072	0.078	0.071	
d_mech	-0.044	0.058	-0.035	0.060	-0.024	0.060	
年度固定効果							
d_1996	0.005	0.075	0.002	0.080	-0.002	0.080	
d_1997	0.050	0.073	-0.052	0.076	-0.052	0.076	
d_1998	0.035	0.076	-0.302	0.080 ***	-0.290	0.080 ***	
d_1999	0.047	0.076	-0.163	0.078 **	-0.175	0.079 **	
d_2000	0.040	0.075	-0.203	0.078 ***	-0.203	0.078 ***	
d_2001	0.050	0.090	-0.167	0.093 *	-0.161	0.093 *	
定数項	-0.291	0.111 ***	0.970	0.252 ***	1.183	0.213 ***	
逆ミルズ比	0.156	0.140					
***:1% sig., **:5% sig., *:10% sig.							

表 5 論文発表とサイエンス・リンケージの関係

		論文発表		
		なし	あり	計
SL	なし	3,842	673	4,515
	あり	387	241	628
計		4,229	914	5,143

num_sl と paper_3 のカテゴリーをそれぞれ“あり”と“なし”に集約しクロス集計した。 χ^2 テストでは独立性が強く棄却される。

参考表 1 “研究開始から特許出願まで (r_a_month)” と “研究開始から利用開始まで (r_u_month)”
月別の度数と累積% (60 ヶ月まで)

月数	r_a_month		r_u_month	
	度数	累積%	度数	累積%
1	88	1.7	0	0.0
2	41	2.6	3	0.1
3	183	6.2	4	0.3
4	44	7.0	11	0.8
5	31	7.7	7	1.1
6	593	19.4	28	2.2
7	10	19.6	20	3.1
8	53	20.6	11	3.6
9	27	21.2	62	6.2
10	36	21.9	17	6.9
11	7	22.0	14	7.5
12	1201	45.8	97	11.6
13	25	46.3	28	12.8
14	78	47.9	13	13.4
15	64	49.1	79	16.7
16	8	49.3	25	17.8
17	16	49.6	10	18.2
18	462	58.7	174	25.6
19	7	58.9	10	26.0
20	12	59.1	23	27.0
21	9	59.3	34	28.4
22	9	59.5	6	28.7
23	5	59.6	8	29.0
24	853	76.5	241	39.2
25	15	76.8	14	39.8
26	13	77.0	9	40.2
27	19	77.4	32	41.6
28	8	77.5	7	41.9
29	11	77.8	6	42.1
30	130	80.3	154	48.7
31	2	80.4	8	49.0
32	8	80.5	16	49.7
33	6	80.7	10	50.1
34	7	80.8	8	50.4
35	2	80.8	6	50.7
36	418	89.1	229	60.4
37	7	89.2	19	61.2
38	12	89.5	9	61.6
39	5	89.6	20	62.5
40	1	89.6	6	62.7
41	1	89.6	5	62.9
42	59	90.8	95	66.9
43	1	90.8	5	67.2
44	1	90.8	1	67.2
45	2	90.9	4	67.4
46	5	91.0	6	67.6
47	0	91.0	4	67.8
48	91	92.8	158	74.5
49	2	92.8	5	74.7
50	4	92.9	8	75.1
51	7	93.0	8	75.4
52	3	93.1	6	75.6
53	2	93.1	3	75.8
54	9	93.3	43	77.6
55	0	93.3	3	77.7
56	1	93.3	3	77.9
57	1	93.3	4	78.0
58	1	93.4	4	78.2
59	0	93.4	1	78.2
60	156	96.5	95	82.3

参考表 2 分析に用いた変数の定義

変数名	内容	変数の種類	単位・カテゴリー	ソース
ラグに関する変数(被説明変数)				
r_a_month	研究開始から特許出願まで	スケール変数	単位は month	サーベイ Q4.4
log_ra_month			上記変数の自然対数 Ln(X+1)	
a_u_month	特許出願から利用開始まで	スケール変数	単位は month	サーベイ Q6.4
log_au_month			上記変数の自然対数 Ln(X+1)	
d_au_month	利用開始までの回答有無	二値変数	基準は”無回答”	サーベイ Q6.4
(説明変数)				
研究プロジェクトの規模				
ra_man_month6	研究開始から特許出願までの投入人・月	順序変数	1:3 人月以下 2:4～6 人月 3:7～12 人月 4:13～24 人月 5:25～48 人月 6:49 人月以上	サーベイ Q4.16
ra_man	研究プロジェクトに関わった推定人数	スケール変数 (合成変数)	r_a_monthとra_man_month6から計算されたおよその人数	サーベイ Q4.16/Q4.4
num_inv	発明者の数	スケール変数	当該特許に記載された発明者の数	特許データベース
num_relpat	関連特許の数	順序変数	1:1 件(当該特許のみ) 2:2～5 件 3:6～10 件 4:11～50 件 5:51～100 件 6:101 件以上	サーベイ Q4.15
d_co_app	共同出願人の有無	二値変数	基準は”単独出願”	特許データベース
研究プロジェクトの性格				
d_basic	研究の範囲:基礎研究	二値変数	非排他的ダミー変数	サーベイ Q4.2
d_applied	研究の範囲:応用研究			
d_poj_tech	研究目的:技術基盤強化	二値変数	基準は”事業に直結する研究”	サーベイ Q4.1
発明の経緯				
d_main_prdc	研究プロジェクトの主目標	二値変数	排他的ダミー変数 基準は”予想もしない副産物”	サーベイ Q4.5
d_sub_prdc	予想可能な副産物			
d_no_rand	研究によらない発明			
d_prev_inv	先行発明への依拠	二値変数	基準は”依拠しない”	サーベイ Q4.7
発明の性格				
d_product_only	発明内容:モノの発明	二値変数	排他的ダミー変数 基準は”モノと方法の両方”	サーベイ Q4.6
d_process_only	発明内容:方法の発明			
技術的な画期性				
d_tri	三極出願かどうか	二値変数	三極出願特許からのサンプル	サンプル属性
d_std_vip	標準・重要特許かどうか	二値変数	標準・重要特許からのサン	サンプル属性

				プル	
	paper_3	論文発表	順序変数 (合成変数)	0:論文として発表していない (技術的水準が高くない) 1:論文として発表していない 2:論文として発表した	サーベイ Q6.9 と Q6.8
企業変数(企業規模)					
	firm_size	企業規模	順序変数	1:従業員 501 人以上の企業 2:従業員 251~500 人の企業 3:従業員 101~250 人の企業 4:従業員 100 人以下の企業	サーベイ Q3.2
技術分野固定効果					
	d_chem	米国分類:化学分野	二値変数	排他的ダミー変数 基準は”その他分野”	特許データベース
	d_comp	米国分類:情報通信分野			
	d_drug	米国分類:医薬・医療分野			
	d_elec	米国分類:電気・電子分野			
	d_mech	米国分類:機械分野			
年度固定効果					
	d_1996	1996 年出願・優先権主張	二値変数	排他的ダミー変数 基準は”1995 年”	特許データベース
	d_1997	1997 年出願・優先権主張			
	d_1998	1998 年出願・優先権主張			
	d_1999	1999 年出願・優先権主張			
	d_2000	2000 年出願・優先権主張			
	d_2001	2001 年出願・優先権主張			
発明の価値の指標					
	num_ipc	IPC サブクラス数	スケール変数	出願時の IPC サブクラスの数	特許データベース
	famsize	ファミリーサイズ	スケール変数	当該特許のファミリーサイズ	特許データベース
	ln_claim	請求項の数	スケール変数	請求項数の自然対数 Ln(X+1)	特許データベース
	d_earl_ex	早期審査請求の有無	二値変数	1 年以内の審査請求の有無	特許データベース
	inv_fc5_or	発明者 5 年前方引用	順序変数	0:被引用なし 1:1 件の被引用 2:2 件以上の被引用	特許データベース
	exm_fc5_or	審査官 5 年前方引用	順序変数	0:被引用なし 1:1 件の被引用 2:2 件以上の被引用	特許データベース
	inv_bc	発明者後方引用数	スケール変数	発明者による先行特許引用数	特許データベース
	num_sl	サイエンスリンケージ数	順序変数	0:科学論文引用なし 1:1 件の科学論文引用 2:2 件以上の科学論文引用	特許データベース
技術分野の競争状態					

itc_total	技術分類サイズ	スケール変数	ITC33 技術分野中当該技術分野・当該年の出願総数	特許データベース
itc_conc	技術分類集中度	スケール変数	ITC33 技術分野中当該技術分野・当該年の上位 3 社集中度	特許データベース
技術の商業化にとって重要な企業戦略				
asset_tech	戦略: 技術開発先行	順序変数	1: まったく重要ではない 2: 重要ではない 3: どちらでもない 4: 重要である 5: 非常に重要である	サーベイ Q6.1
asset_lead	戦略: 市場投入先行			
asset_manu	戦略: 製造能力			
asset_sales	戦略: 販売能力			
asset_enfor	戦略: 特許権エンフォース			
asset_know	戦略: 企業秘密			
asset_cmplx	戦略: 複雑性			
asset_coop	戦略: 外部企業協力			

参考表 3 説明変数の記述統計

変数	度数 有効回答	欠損値	最小値	最大値	平均値	中央値	最頻値	標準偏差	歪度	尖度
ra_man_month6	5,114	136	1.00	6.00	3.21	3.00	1.00	1.695	0.17	-1.21
ra_man	4,981	269	0.00	49.00	0.99	0.54	0.33	2.111	12.63	240.52
num_inv	5,250	0	1.00	21.00	2.60	2.00	1.00	1.721	1.77	6.18
num_relpat	5,117	133	1.00	6.00	2.31	2.00	2.00	1.087	1.06	1.05
d_co_app	5,250	0	0.00	1.00	0.15	0.00	0.00	0.358	1.95	1.80
d_basic	5,221	29	0.00	1.00	0.21	0.00	0.00	0.408	1.41	0.00
d_applied	5,221	29	0.00	1.00	0.37	0.00	0.00	0.484	0.52	-1.73
d_poj_tech	5,250	0	0.00	1.00	0.09	0.00	0.00	0.292	2.78	5.74
d_main_prdc	5,250	0	0.00	1.00	0.45	0.00	0.00	0.498	0.20	-1.96
d_sub_prdc	5,250	0	0.00	1.00	0.23	0.00	0.00	0.423	1.26	-0.42
d_no_rand	5,250	0	0.00	1.00	0.14	0.00	0.00	0.35	2.05	2.19
d_prev_inv	5,250	0	0.00	1.00	0.49	0.00	0.00	0.5	0.03	-2.00
d_product_only	5,250	0	0.00	1.00	0.44	0.00	0.00	0.496	0.25	-1.94
d_process_only	5,250	0	0.00	1.00	0.34	0.00	0.00	0.473	0.69	-1.53
d_tri	5,250	0	0.00	1.00	0.70	1.00	1.00	0.46	-0.86	-1.27
d_std_vip	5,250	0	0.00	1.00	0.02	0.00	0.00	0.131	7.40	52.76
paper_3	5,143	107	0.00	2.00	1.13	1.00	1.00	0.457	0.49	1.14
firm_size	4,948	302	1.00	4.00	3.70	4.00	4.00	0.784	-2.63	5.63
d_chem	5,250	0	0.00	1.00	0.18	0.00	0.00	0.382	1.70	0.87
d_comp	5,250	0	0.00	1.00	0.14	0.00	0.00	0.347	2.08	2.31
d_drug	5,250	0	0.00	1.00	0.07	0.00	0.00	0.251	3.44	9.83
d_elec	5,250	0	0.00	1.00	0.17	0.00	0.00	0.38	1.71	0.94
d_mech	5,250	0	0.00	1.00	0.19	0.00	0.00	0.39	1.60	0.57
d_1996	5,250	0	0.00	1.00	0.15	0.00	0.00	0.359	1.94	1.76
d_1997	5,250	0	0.00	1.00	0.17	0.00	0.00	0.377	1.74	1.03
d_1998	5,250	0	0.00	1.00	0.15	0.00	0.00	0.354	1.99	1.97
d_1999	5,250	0	0.00	1.00	0.16	0.00	0.00	0.369	1.83	1.35
d_2000	5,250	0	0.00	1.00	0.17	0.00	0.00	0.376	1.75	1.07
d_2001	5,250	0	0.00	1.00	0.09	0.00	0.00	0.279	2.97	6.85
num_IPC	5,250	0	1.00	68.00	3.84	3.00	2.00	3.22	5.12	61.93
famsize	5,197	53	1.00	29.00	3.95	4.00	1.00	2.613	1.77	8.70
Ln_claim	5,250	0	0.00	5.17	1.68	1.79	1.10	0.843	-0.09	-0.01
d_earl_ex	5,250	0	0.00	1.00	0.13	0.00	0.00	0.333	2.24	3.02
inv_fc5_or	5,250	0	0.00	2.00	0.49	0.00	0.00	0.776	1.17	-0.33
exm_fc5_or	5,250	0	0.00	2.00	0.31	0.00	0.00	0.624	1.84	2.01
inv_bc	5,250	0	0.00	100.00	1.60	1.00	0.00	3.668	10.91	214.21
num_SL	5,250	0	0.00	2.00	0.19	0.00	0.00	0.544	2.71	5.83
itc_total	5,234	16	175	55142	23203	14419	54109	16979	0.63	-1.15
itc_conc	5,234	16	0.03	0.49	0.13	0.14	0.14	0.056	0.72	1.55
asset_tech	5,017	233	1.00	5.00	4.03	4.00	4.00	0.869	-1.14	1.83
asset_lead	5,003	247	1.00	5.00	3.95	4.00	4.00	0.868	-1.01	1.51
asset_manu	5,000	250	1.00	5.00	3.81	4.00	4.00	0.931	-0.88	0.89
asset_sales	4,985	265	1.00	5.00	3.58	4.00	4.00	0.966	-0.62	0.30
asset_enfor	4,970	280	1.00	5.00	3.47	4.00	4.00	0.863	-0.59	0.48
asset_know	4,976	274	1.00	5.00	3.58	4.00	4.00	0.961	-0.62	0.26
asset_cmplx	4,951	299	1.00	5.00	3.00	3.00	3.00	0.918	-0.25	0.08
asset_coop	4,974	276	1.00	5.00	3.19	3.00	3.00	1.003	-0.39	-0.14